

SEQUENCE LISTING

<110> Genencor International, Inc.

Jones, Brian E.

Grant, William D.

Heaphy, Shaun

Grant, Susan

<120> Novel Bacillus BagCel Cellulase

<130> GC798-2-PCT

<140> PCT/US2004/013175

<141> 2004-04-28

<150> US 60/467,255

<151> 2003-04-30

<160> 3

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 4205

<212> DNA

<213> Bacillus sp.

<220>

<221> misc_feature

<222> (1)...(4205)

<223> isolated from environmental water sample from
Sonachi Lake, Kenya

<400> 1

tctcatttag	tttgactgaa	atacctactg	tgccacaaaag	taaagttaaa	ctgacgaatg	60
tggAACGAAT	CACTTAATGG	TTCTAGCATA	GATAACGAAA	GATGAACACG	TTCAAAGTTC	120
gccACTCTT	TGAAAGAGGG	TGAACTTTT	TTGTGACAAG	AAAGGGTGT	AAATGAAGAT	180
caaAGCTAAA	CAAGATGAGT	AACGTTCTT	TTCTTTTTA	TAGAGTGAGT	TAGTATATGA	240
gtCCCTTATA	AATTCTAGA	CTGTTATTT	AAATAATTGA	ATGACTCAGT	CACCATTAAG	300
ttttcaacac	CCATAAGCGA	CGTTGAAGA	TCTAATGATG	CGAGAGGTT	TATCACTTTG	360
gagCgGAAGA	TCACTGTAGG	ACTCGTTTA	TATGGTGAAC	TTGGTGTAT	TGTGTATT	420
aaAGAAAGGG	AAACGAAAAAA	AATGGTTAAA	TTAGAAAGAG	GCTATTACAG	AGAGGAGAAC	480
aaacaatgaa	CGTAACACTT	GAAGTGAACAT	ACTGCACGAC	TAAGGTATT	CGAACAACCT	540
ttcattcaga	AGGTATGGAG	GCGAAAAAAG	CAATTACCAT	CGCAGAAAGAT	TTTCAGCGGA	600
caggacggat	AAAACAGATC	GTCTTAGAG	ATGAGCGTGA	TAGTCCGTGG	ACGTAAAAG	660
aacttAAAAG	ATTTTAAAGA	GAGATTAAAA	CGGAGCCGCA	TCACTCTCT	GTGTATT	720
atgggggatt	TGATTGGAG	ACACAAACGAT	CTGGTCTTGG	GTGTGATT	TTATGAACAA	780
aatgacacgt	CTTATCGGGT	GAGAAGAAAC	GCTACCGTGG	CGTCATTGAC	ATCGAATAAC	840
gaagcagaat	ATGCCGCTT	ACATTAGGA	CTTAAAGAAC	TTGAAGGGAT	CGGTGCGCAT	900
catctaccta	TCACTATT	CGGTGATTCT	CAAGTTGTGA	TCAATCAGTT	AAAAGGAGAA	960
tgggcgtgt	TGGAGGAGGT	GTAAATAAAA	TGGGCTGACC	GTATTGATCA	GCATTAGCT	1020
aaattaggca	TGACCGCTAC	TTATAAGTTA	ATCCCCCGTA	AAGAAAACCG	TGAAGCAGAT	1080
caactggcta	CACAAAGCGTT	AAACGGGCAA	GAATTATAA	GTCAACGTGA	TGTCACTGAG	1140
cgtggcgcag	ATTAGTCCTGC	ACCCGCATAA	AAGTTAACGT	ATATAGAACGT	GGATGGGGAT	1200

taaaggaacg	tcattcaactc	taagcaagcg	ttgcgcacagc	aaaaaaagaaaa	catataagg	1260
ttttctgagc	tactatctat	acaatagcc	aagtggcagt	taagcttta	cctcatcaag	1320
tttttgacta	ccagtctcc	actcctactt	tcacctatat	aaattggttc	ctttttgtt	1380
aataatcaact	aattttgacg	gtatTTTTA	atagaatat	atgctagatt	ataaaactagt	1440
aacgatgtag	aagggtgtga	ttgaccat	aagaagactc	tttcaaacct	ggttagtatcg	1500
cattaaaaaa	tttggaaaggt	ggagaggaca	catgggtat	accaaagcga	agtgtacgtt	1560
gaaaaaaaaact	gtcttggttt	gtttaattct	ctgttaagt	gtgtcaatgt	ttgttccaat	1620
gacatcagct	gaagatgtca	cttcgtcaca	gttggatatt	cactcctatg	tagctgacat	1680
gcagcctggc	ttggatttag	gaaatacg	tgacgcgtt	ggagatgatg	aaacagcgtg	1740
ggggaaatcct	cgtgtacaaa	gagagttaat	aaaaacgatt	gctgatgaag	ggtataaaag	1800
cattcgtatc	ccagtgcacat	ggcaaaaatca	aatgggtgg	tctccagatt	atacgataaa	1860
tgaagattat	atcaatcggg	tggagcaagc	gatagattgg	gcgttggagg	aagacttata	1920
tgtatgtta	aatgtgcac	atgactcatg	gctgtggatg	tatgatatgg	aacataaacta	1980
tgtatgggtc	atggcaagat	atacagctat	ttgggaacaa	ttgtcgaaa	aattcaaaag	2040
ccactccat	aagttgtatgt	ttgagagtgt	caatgagcct	aggtttacgc	aggagtgggg	2100
agagattcaa	gaaaatcatc	atgcttactt	agaagattta	aataagacgt	tctattat	2160
tgtcagagag	tcaggaggca	ataatgtga	gcgccttta	gtattgccta	cgatagaaac	2220
agccacgtct	caggatttac	tagatcgctt	gtatcaaaca	atggaagact	tggatgatcc	2280
ttatTTTatt	gccacggtgc	attattatgg	cttctggcca	tttagtgtca	atatacgagg	2340
gtacactcat	tttgaacagg	aaacacaaca	agatattata	gacaccttgc	accgtgttca	2400
taacacattt	acagcgctg	gtgtcccagt	tgtattagc	gaattcggtt	tgttaggctt	2460
tgacaaaagt	acggatgtga	ttcagcaagg	ggagaaaatta	aagtttttg	agtttctcat	2520
ccatcatctc	aatgaacgtg	atataaccca	tatgttatgg	gataacggcc	agcattaaa	2580
tcgagaaaact	tatgcattgt	atgatcaaga	atttcatgac	atattaaaag	cgagtggga	2640
ggggcgttct	gctacagcag	agtctaattt	gattcatgt	aaggacggaa	agccaattag	2700
agatcaagat	atacagctt	acttaaacgg	aatgagcta	acagccttac	aggcagggga	2760
ggaatcgctt	gttcttaggag	aggattatga	actagcagga	ggcgtattaa	cgctaaaagc	2820
ggacaccctc	acaagactaa	ttacccctgg	tcaattagga	accaatgcag	tcatcacagc	2880
acaatTTAAT	tctggagcag	actggcg	tcaattacag	aatgtggacg	tgccaacgg	2940
cgaaaataca	gatggctcaa	catggcattt	tgcgatccct	acccatttta	atggtgatag	3000
tcttgcacg	atggaaagctg	tttatgc	cgagaatat	gttggccgc	aagattggac	3060
gtcatttaaa	gaatttggcg	aggcgttt	tcctaattac	gcccacagggg	aaattattat	3120
atcagaagcc	ttctttaacg	cggtacggg	tgatgat	catttaacat	ttcattttt	3180
gagcggagag	acggtggaaat	ataccctacg	taaaaatggc	aattatgttc	aaggtagacg	3240
gtaacatgat	tttaattat	agataaacc	gcctacctat	cgttttgg	agaaggcaaa	3300
cgaatctcat	tttaccaaca	ccgtcttta	gaactttaga	agtgcgg	tttttaaga	3360
catgaggaga	agcaatctc	tatcaacagt	caccaattt	tattcaggag	gtgtcaagtt	3420
atctaacgtt	ctatgaatgc	atatagtctc	tgacgaataa	acatagttaa	aaagaagtga	3480
gcctagtccc	cgaggggaag	gggataatgc	caacgtatt	gattaaagta	ccttcttgat	3540
aaaaagaaaag	gttttcaag	agatggaaat	gggctcg	tttatactt	aattacgcct	3600
tggAACGTCA	ttttggcggt	gttatttagt	aagatggctg	acatcataaa	agaggagtgg	3660
gttcgtatggc	tttaattcaa	ttaagctt	aatcagcagc	attaatgtt	caaacctctg	3720
tcaatgttt	attaccgg	ggaatgaatg	ctgttagattt	tacaccaagt	gatgatTTT	3780
cttatgttac	tgaccctt	cctgtcctat	atcttttgc	tggtgcaact	gatgattatt	3840
cagcatggct	acgtctgtcc	tctattgaac	gatatgctg	agaaaaaaa	ttggcggtcg	3900
tcatgccaaa	tgctgat	agtgcgtata	cggtatgtt	acatggccat	cgttactgga	3960
cgtatattag	taaggtgctg	cctgagtt	tgagagcaac	tttcctatt	tctcagcacc	4020
gtgaagacac	ctttgcagct	ggctgtct	tggaggata	cggggctt	aaattggcgc	4080
tgccgcaacc	ggaacgc	gtcgcagctg	tgtcattatc	tggtgca	gatatgagag	4140
aagcaagtca	accagactcc	ctatttgtaa	atgccttgg	tgaaggacg	aaaatcg	4200
ggaca						4205

<210> 2

<211> 1713

<212> DNA

<213> Bacillus sp.

<220>
 <221> misc_feature
 <222> (1)...(1713)
 <223> isolated from environmental water sample from
 Sonachi Lake, Kenya

<400> 2

atgggttata	ccaaagcgaa	gtgtacgttg	aaaaaaaactg	tottgtttgg	ttaattctc	60
tgtttaagt	tgtcaatgt	tgtccaatg	acatca	agatgtcac	ttcgtcacag	120
ttggatattc	actcctatgt	agctgacatg	cagcctggct	ggaatttagg	aaatacgtt	180
gacgctgtt	gagatgtga	aacagcgtgg	ggaaatc	gtgtaca	agagtaata	240
aaaacgattt	ctgtatgaagg	gtataaaagc	attcgat	cagtgcacatg	gcaaata	300
atgggtgg	ctccagatta	tacgataaaat	gaagattata	tcaatcgggt	ggagcaagcg	360
atagattgg	cggtggagga	agacttat	gtgtat	atgtgc	tgactcatgg	420
ctgtggatgt	atgatatgg	acataactat	gatgagg	tggcaagata	tacagctatt	480
tggaaacaat	tgtcgaaaaa	attcaaaagc	cactcccata	agttgatgtt	tgagagtgtc	540
aatgagccta	ggtttacgca	ggagtggg	gagattcaag	aaaatcatca	tgcttactta	600
gaagatttaa	ataagacgtt	ctattatatt	gtcagagat	caggaggcaa	taatgtggag	660
cgcctttag	tattgcctac	gatagaaaca	gccacgtc	aggattact	agatcgctt	720
tatcaaacaa	tggaagactt	ggatgatcct	tatthaattt	ccacgggtca	ttattatggc	780
ttctggccat	ttagtgtcaa	tatagcaggg	tacactcatt	ttgaacagga	aacacaacaa	840
gatattatag	acaccttga	ccgtgttcat	aacacattt	cagcgcgtgg	tgtcccagtt	900
gtattaggcg	aattcggtt	gttaggctt	gacaaaagta	cggatgtgat	tcagcaaggg	960
gagaaattaa	agtttttga	gtttctcatc	catcatctca	atgaacgtga	tataaccat	1020
atgttatgg	ataacggcca	gcatttaat	cgagaaactt	atgc	tgatcaagaa	1080
tttcatgaca	tataaaagc	gagttggag	ggcggttct	ctacagcaga	gtctaattt	1140
attcatgtga	aggacggaaa	gccaaattt	gatcaagata	ta	cttaaacgg	1200
aatgagctaa	cagccttaca	ggcaggggag	gaatcgctt	ttctaggaga	ggattatgaa	1260
ctagcaggag	gcgtttaac	gctaaaagcg	gacacc	caagacta	taccctggt	1320
caatttagaa	ccaatgcagt	catcacagca	caattt	ctggagcaga	ctggcg	1380
caattacaga	atgtggacgt	gccaacgg	gaaaatacag	atggctca	atggcattt	1440
gcgatcccta	cccat	ttgtat	cttgc	tggaa	ctgtcaa	1500
ggagaatatg	ctggccgca	agattggacg	tcattt	aa	atttggcga	1560
cctaattacg	ccacaggg	aattattata	tcaga	acct	tcttaacgc	1620
gatgatatcc	at	ttaacatt	tcattt	ggagaga	cggtgaaata	1680
aaaaatggca	attatgttca	aggtagacgg	taa			1713

<210> 3
 <211> 570
 <212> PRT
 <213> Bacillus sp.

<220>
 <221> VARIANT
 <222> (1)...(570)
 <223> isolated from environmental water sample from
 Sonachi Lake, Kenya

<400> 3

Met	Gly	Tyr	Thr	Lys	Ala	Lys	Cys	Thr	Leu	Lys	Lys	Thr	Val	Leu	Phe
1					5				10				15		
Gly	Leu	Ile	Leu	Cys	Leu	Ser	Val	Ser	Met	Phe	Val	Pro	Met	Thr	Ser
									20			25		30	
Ala	Glu	Asp	Val	Thr	Ser	Ser	Gln	Leu	Asp	Ile	His	Ser	Tyr	Val	Ala
									35			40		45	

Asp Met Gln Pro Gly Trp Asn Leu Gly Asn Thr Phe Asp Ala Val Gly
50 55 60
Asp Asp Glu Thr Ala Trp Gly Asn Pro Arg Val Thr Arg Glu Leu Ile
65 70 75 80
Lys Thr Ile Ala Asp Glu Gly Tyr Lys Ser Ile Arg Ile Pro Val Thr
85 90 95
Trp Gln Asn Gln Met Gly Gly Ser Pro Asp Tyr Thr Ile Asn Glu Asp
100 105 110
Tyr Ile Asn Arg Val Glu Gln Ala Ile Asp Trp Ala Leu Glu Glu Asp
115 120 125
Leu Tyr Val Met Leu Asn Val His His Asp Ser Trp Leu Trp Met Tyr
130 135 140
Asp Met Glu His Asn Tyr Asp Glu Val Met Ala Arg Tyr Thr Ala Ile
145 150 155 160
Trp Glu Gln Leu Ser Glu Lys Phe Lys Ser His Ser His Lys Leu Met
165 170 175
Phe Glu Ser Val Asn Glu Pro Arg Phe Thr Gln Glu Trp Gly Glu Ile
180 185 190
Gln Glu Asn His His Ala Tyr Leu Glu Asp Leu Asn Lys Thr Phe Tyr
195 200 205
Tyr Ile Val Arg Glu Ser Gly Gly Asn Asn Val Glu Arg Pro Leu Val
210 215 220
Leu Pro Thr Ile Glu Thr Ala Thr Ser Gln Asp Leu Leu Asp Arg Leu
225 230 235 240
Tyr Gln Thr Met Glu Asp Leu Asp Asp Pro Tyr Leu Ile Ala Thr Val
245 250 255
His Tyr Tyr Gly Phe Trp Pro Phe Ser Val Asn Ile Ala Gly Tyr Thr
260 265 270
His Phe Glu Gln Glu Thr Gln Gln Asp Ile Ile Asp Thr Phe Asp Arg
275 280 285
Val His Asn Thr Phe Thr Ala Arg Gly Val Pro Val Val Leu Gly Glu
290 295 300
Phe Gly Leu Leu Gly Phe Asp Lys Ser Thr Asp Val Ile Gln Gln Gly
305 310 315 320
Glu Lys Leu Lys Phe Phe Glu Phe Leu Ile His His Leu Asn Glu Arg
325 330 335
Asp Ile Thr His Met Leu Trp Asp Asn Gly Gln His Leu Asn Arg Glu
340 345 350
Thr Tyr Ala Trp Tyr Asp Gln Glu Phe His Asp Ile Leu Lys Ala Ser
355 360 365
Trp Glu Gly Arg Ser Ala Thr Ala Glu Ser Asn Leu Ile His Val Lys
370 375 380
Asp Gly Lys Pro Ile Arg Asp Gln Asp Ile Gln Leu Tyr Leu Asn Gly
385 390 395 400
Asn Glu Leu Thr Ala Leu Gln Ala Gly Glu Glu Ser Leu Val Leu Gly
405 410 415
Glu Asp Tyr Glu Leu Ala Gly Gly Val Leu Thr Leu Lys Ala Asp Thr
420 425 430
Leu Thr Arg Leu Ile Thr Pro Gly Gln Leu Gly Thr Asn Ala Val Ile
435 440 445
Thr Ala Gln Phe Asn Ser Gly Ala Asp Trp Arg Phe Gln Leu Gln Asn
450 455 460
Val Asp Val Pro Thr Val Glu Asn Thr Asp Gly Ser Thr Trp His Phe
465 470 475 480
Ala Ile Pro Thr His Phe Asn Gly Asp Ser Leu Ala Thr Met Glu Ala
485 490 495
Val Tyr Ala Asn Gly Glu Tyr Ala Gly Pro Gln Asp Trp Thr Ser Phe

500 505 510
Lys Glu Phe Gly Glu Ala Phe Ser Pro Asn Tyr Ala Thr Gly Glu Ile
515 520 525
Ile Ile Ser Glu Ala Phe Phe Asn Ala Val Arg Asp Asp Asp Ile His
530 535 540
Leu Thr Phe His Phe Trp Ser Gly Glu Thr Val Glu Tyr Thr Leu Arg
545 550 555 560
Lys Asn Gly Asn Tyr Val Gln Gly Arg Arg
565 570